

Ocean color chlorophyll algorithm round robin for the Chesapeake Bay

P. Jeremy Werdell
NASA Ocean Biology Processing Group
Science Systems and Applications, Inc.

15 June 2006

General Purpose

Satellite-derived ocean color data products provide the scientific community a means of studying the marine biosphere on temporal and regional scales unattainable by conventional *in situ* methods. As such, the Chesapeake Bay Program has interest in using SeaWiFS and MODIS chlorophyll *a*, C_a , data products to facilitate their regional Chesapeake Bay monitoring activities. In collaboration with the NASA Ocean Biology Processing Group (OBPG) and the University of Maryland, the Chesapeake Bay Program initiated an algorithm round robin to evaluate the performance of currently available remote sensing C_a products for the Bay. The OBPG compiled nine candidate algorithms for this activity and accepted responsibility for satellite data processing and preliminary data analysis.

With regards to the latter, the candidate remotely sensed products are initially evaluated via direct comparison with *in situ* C_a measurements. In particular, coincident satellite and *in situ* values are statistically compared and long-term time-series and data distributions (e.g., histograms) are evaluated. This document provides a brief description of the candidate algorithms and outlines the relevant processing and analysis methods applied by the OBPG. Preliminary results for SeaWiFS are available online at <http://seabass.gsfc.nasa.gov/eval/cbp_eval.cgi>.

Ocean Color Chlorophyll Algorithms

Nine C_a algorithms were considered in this activity. Several were designed for global use, for example, the operational algorithms of SeaWiFS, MODIS, and VIIRS. Others were specifically created for the Chesapeake Bay. For this analysis, all are tuned to the specific spectral resolution of SeaWiFS.

1. OC4 version 4 (OC4v4; SeaWiFS operational algorithm; O'Reilly et al. 2000)
2. OC4 version 5 (OC4v5)
3. OC3 version 5 (OC3v5; MODIS operational algorithm)
4. OC2 version 5 (OC2v5)
5. OC3-CB (provided by Old Dominion University)
6. Clark (provided by NOAA)
7. GSM01 (Maritorena et al. 2002)
8. GSM01-CB (Magnuson et al. 2004)
9. Carder (VIIRS operational algorithm; Carder et al. 1999)

The first six algorithms were empirically derived through the statistical correlation of coincident *in situ* radiometric and C_a measurements (Table 1). Following O'Reilly et al. (2000), the OC version 5 algorithms were recently derived using the NASA bio-Optical Marine Algorithm Data set (NOMAD; Werdell and Bailey 2005). In contrast, while algorithms 7 – 9 include empirical components, their C_a solution requires the inversion of a simplified form of the radiative transfer equation. Refer to the provided references and Appendix A of Signorini et al. (2005) for a detailed description of the latter.

Data Stratification

Given its geographic variability, the Bay was regionally stratified for this analysis following Magnuson et al. (2004) (Figure 1):

1. Upper Bay $> 38.6^\circ \text{ N}$
2. Mid Bay $37.6 - 38.6^\circ \text{ N}$
3. Lower Bay $< 37.6^\circ \text{ N}$

Similarly, the data were temporally stratified by Season (modified for Leap Years):

1. Spring day numbers 80 to 172
2. Summer day numbers 173 to 266
3. Fall day numbers 267 to 355
4. Winter day numbers 1 to 79 and 356 to 365

In Situ Data

Acquisition. We acquired approximately 15,750 discrete fluorometric C_a samples collected by participants in the Chesapeake Bay Program (Chesapeake Bay Program 1993). These data were supplemented with approximately 2,300 independent discrete fluorometric C_a samples collected by Dr. L.W. Harding of the University of Maryland (Figure 2). With regards to the latter, high performance liquid chromatography samples are also available in lieu of the fluorometric values for subsequent analyses. Given the general turbidity of the Bay – and resulting shallow optical depths – only near-surface samples (≤ 1 meter) were considered. Replicate samples were averaged. Additional details regarding the treatment of the *in situ* data are available in Werdell and Bailey (2005).

Data distributions and time-series. Regional histograms and time-series were generated for comparison with the satellite C_a data. Only retrievals within $0 \leq C_a \leq 100 \text{ mg m}^{-3}$ were considered, as, in general, this is the effective operational range of the algorithms under consideration (O'Reilly et al. 2000). Only 24 of the 18,000 available stations had values greater than 100 mg m^{-3} . On the contrary, 52 and 8 stations had C_a less than 1.0 and 0.3 mg m^{-3} , respectively. The lowest C_a recorded was 0.17 mg m^{-3} .

In generating the histograms, all stations were considered. The histograms are presented in lognormal space (Campbell 1995), with a bin size of 0.05 (Figure 3). The sample median and mode are also presented. For the time-series, we calculated the monthly mean, median, and standard deviation (on a year-by-year basis) of all stations for each geographic region (Figures 4 and 5). The latter improved visual clarity in the time-series plots and, in general, minimized the overall impact of discrete anomalous stations.

Satellite Data

Acquisition. Approximately 5,000 extracted SeaWiFS Merged Local Area Coverage (MLAC) Level-1A files were acquired from the OBP using the OceanColor Web (<http://oceancolor.gsfc.nasa.gov>). We considered the full, continuous SeaWiFS time series (September 1997 through mid-March 2006). While the original region of interest was the Chesapeake Bay, the extracts encompassed the Mid-Atlantic Bight ($30-45^\circ \text{ N}$, $68-82^\circ \text{ W}$) to facilitate subsequent offshore analyses.

Processing. The Level-1A files were processed to Level-2 using MSL12 version 5.4.1. The operational SeaWiFS atmospheric correction scheme was applied (Gordon and Wang 1994), including the corrections for non-zero near infrared water-leaving radiances (Patt et al. 2003) and bi-directional reflectance (Morel et al. 2002, Franz et al. 2005). Excluding the Clark algorithm, a correction for out-of-band radiance was also required (Wang et al. 2001). We applied the operational SeaWiFS pixel-masking scheme (Patt et al. 2003), with the exception that pixels with stray light contamination were retained. The resultant Level-2 files included the nine candidate C_a data products, plus spectral normalized water-leaving radiances, the downwelling diffuse attenuation coefficient at 490-nm, and the aerosol optical thickness at 865-nm.

File exclusion. We first applied an automated data screening mechanism. Scenes were eliminated when: (a) any pixel within the Bay had a satellite zenith angle greater than 60° (as determined by the SeaWiFS HISATZEN flag); or (b) less than 25% of all Bay pixels were cloud-free (as determined by the SeaWiFS CLDICE flag). We then visually inspected all scenes to determine their suitability for this analysis – in general, removing scenes with large off-Bay cloud fronts and high satellite zenith angles.

Remapping. The remaining SeaWiFS scenes were extracted to a Bay-sized rectangle (36.5-39.5° N, 75-77.5° W), remapped to a fixed equidistant cylindrical projection, and saved as generic HDF files. These files are used the subsequent statistical analyses. True color and OC4v5 C_a images are available online at <http://seabass.gsfc.nasa.gov/eval/cbp_scenes.cgi> for community perusal (Figure 6).

Data distributions and time-series. Regional histograms and time-series were generated for comparison with the *in situ* C_a data. Only retrievals within $0 \leq C_a \leq 100 \text{ mg m}^{-3}$ were considered, as, in general, this is the effective operational range of the algorithms under consideration (O'Reilly et al. 2000). Further, when spatially stratifying the data, points were excluded when sample sizes for a given region were less than 200 valid marine pixels. In generating the histograms, all valid marine pixels for each scene were considered. The histograms are presented in lognormal space (Campbell 1995), with a bin size of 0.05 (Figure 4). The sample median and mode are also presented. For the time-series, we calculated the monthly mean and standard deviation (on a year-by-year basis) of all valid marine pixels for all available scenes for each geographic region (Figures 4 and 5).

Satellite-to-*in situ* Matchups

For each satellite C_a data product, coincident SeaWiFS and *in situ* values were statistically compared (“matched-up”), as described in Bailey and Werdell (2006). In brief, data from both sources were processed as described above. We defined the temporal threshold for coincidence as ± 3 hours. SeaWiFS C_a values were determined as the filtered median of all valid pixels within a 5x5 pixel box centered on the *in situ* target. Satellite pixel exclusion criteria and additional satellite homogeneity tests are graphically described in Figure 7 (Figure 1 from Bailey and Werdell (2006)). Our analysis output includes SeaWiFS-to-*in situ* scatter plots with relevant statistics (Figure 8).

Contact Information

Jeremy Werdell	jeremy.werdell@gsfc.nasa.gov
Larry Harding	larry@hpl.umces.edu
Mike Mallonee	mallonee@hpl.umces.edu
Chuck McClain	charles.r.mcclain@nasa.gov
Gene Feldman	gene.c.feldman@nasa.gov

References

- Bailey, S.W. and P.J. Werdell (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, 102, 12-23.
- Campbell, J.W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research*, 100, 13237-13254.
- Carder, K.L., Lee, Z.P., Hawes, S.K., and Kamykowski, D. (1999). Semianalytic Moderate-Resolution Imaging Spectrometer algorithms for chlorophyll a and absorption with bio-optical domains based on nitrate-depletion temperatures. *Journal of Geophysical Research*, 104, 5403-5421.
- Chesapeake Bay Program (1993). Guide to using Chesapeake Bay Program Water Quality Monitoring Data. CBP/TRS 78/92, Chesapeake Bay Program, Annapolis, Maryland, 127 pp.
- Franz, B.A., Werdell, P.J., Meister, G., Bailey, S.W., Eplee Jr., R.E., Feldman, G.C., Kwiatkowska, E., McClain, C.R., Patt, F.S., and Thomas, D. (2005). The continuity of ocean color measurements from SeaWiFS to MODIS. *Proceedings of SPIE Vol. 5882*, doi: 10.1117/12.620069.
- Gordon, H.R., and Wang, M.H. (1994). Retrieval of water-leaving radiance and aerosol optical-thickness over the oceans with SeaWiFS – a preliminary algorithm. *Applied Optics*, 33, 443-452.
- Magnuson, A., Harding, L.W., Mallonee, M.E., and Adolf, J.E. (2004). Bio-optical model for Chesapeake Bay and the Middle Atlantic Bight. *Estuarine, Coastal, and Shelf Science*, 61, 403-424.
- Maritorena, S., Siegel, D.A., and Peterson, A.R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Applied Optics*, 41, 2705-2714.
- Morel, A., Antoine, D., and Gentili, B. (2002). Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. *Applied Optics*, 41, 6289-6306.
- O'Reilly, J.E. and 24 co-authors (2000). SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3. NASA/TM-2000-206892, Vol. 11, NASA Goddard Space Flight Center, Greenbelt, Maryland, 49 pp.
- Patt, F.S. and 17 co-authors (2003). Algorithm Updates for the Fourth SeaWiFS Reprocessing. NASA/TM-2003-206892, Vol. 22, NASA Goddard Space Flight Center, Greenbelt, Maryland, 74 pp.
- Signorini, S.R., McClain, C.R., Mannino, A. and Bailey, S.W. (2005). Report on Ocean Color and Carbon Study for the South Atlantic Bight and Chesapeake Bay Regions. NASA/TM-2005-212787, NASA Goddard Space Flight Center, Greenbelt, Maryland, 45 pp.
- Wang, M., Franz, B.A., Barnes, R.A., and McClain, C.R. (2001). Effects of spectral bandpass on SeaWiFS-derived near-surface optical properties of the ocean. *Applied Optics*, 40, 343-348.
- Werdell, P.J. and Bailey, S.W. (2005). An improved in situ bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, 98, 122-140.

Tables

Table 1. The form of the six empirical SeaWiFS C_a algorithms under consideration. All require the input SeaWiFS radiances to be corrected for the spectral out-of-band response (Wang et al. 2001), with the exception of the Clark algorithm. The input radiances are in the form of either remote-sensing reflectance, $R_{rs}(\lambda)$, or normalized water-leaving radiance, $L_{wn}(\lambda)$. The OC version 5 algorithms are courtesy of J.E. O'Reilly (NOAA) and are not to be redistributed or publicly released without his or OBPG consent.

Name	Expression	Maximum Band Ratio
OC4v4	$\log_{10} \{C_a\} = 0.366 - 3.067 r + 1.930 r^2 + 0.649 r^3 - 1.532 r^4$	$r = \log_{10} \{(R_{rs443} > R_{rs490} > R_{rs510}) / R_{rs555}\}$
OC4v5	$\log_{10} \{C_a\} = 0.308 - 3.088 r + 3.044 r^2 - 1.201 r^3 - 0.799 r^4$	$r = \log_{10} \{(R_{rs443} > R_{rs490} > R_{rs510}) / R_{rs555}\}$
OC3v5	$\log_{10} \{C_a\} = 0.241 - 2.477 r + 1.530 r^2 + 0.106 r^3 - 1.108 r^4$	$r = \log_{10} \{(R_{rs443} > R_{rs490}) / R_{rs555}\}$
OC2v5	$\log_{10} \{C_a\} = 0.237 - 2.454 r + 1.711 r^2 - 0.340 r^3 - 2.788 r^4$	$r = \log_{10} \{R_{rs490} / R_{rs555}\}$
OC3-CB	$\log_{10} \{C_a\} = -0.115 - 3.678 r$	$r = \log_{10} \{(R_{rs443} > R_{rs490}) / R_{rs555}\}$
Clark	$\log_{10} \{C_a\} = 0.212 - 1.778 r + 0.229 r^2 + 0.423 r^3 + 1.826 r^4 - 2.550 r^5$	$r = \log_{10} \{L_{wn443} / L_{wn555}\}$

Figures

Figure 1. Graphical representation of the Upper (red), Mid (blue), and Lower (green) Bay regions used in this analysis, following the definitions provided in Magnuson et al. (2004).

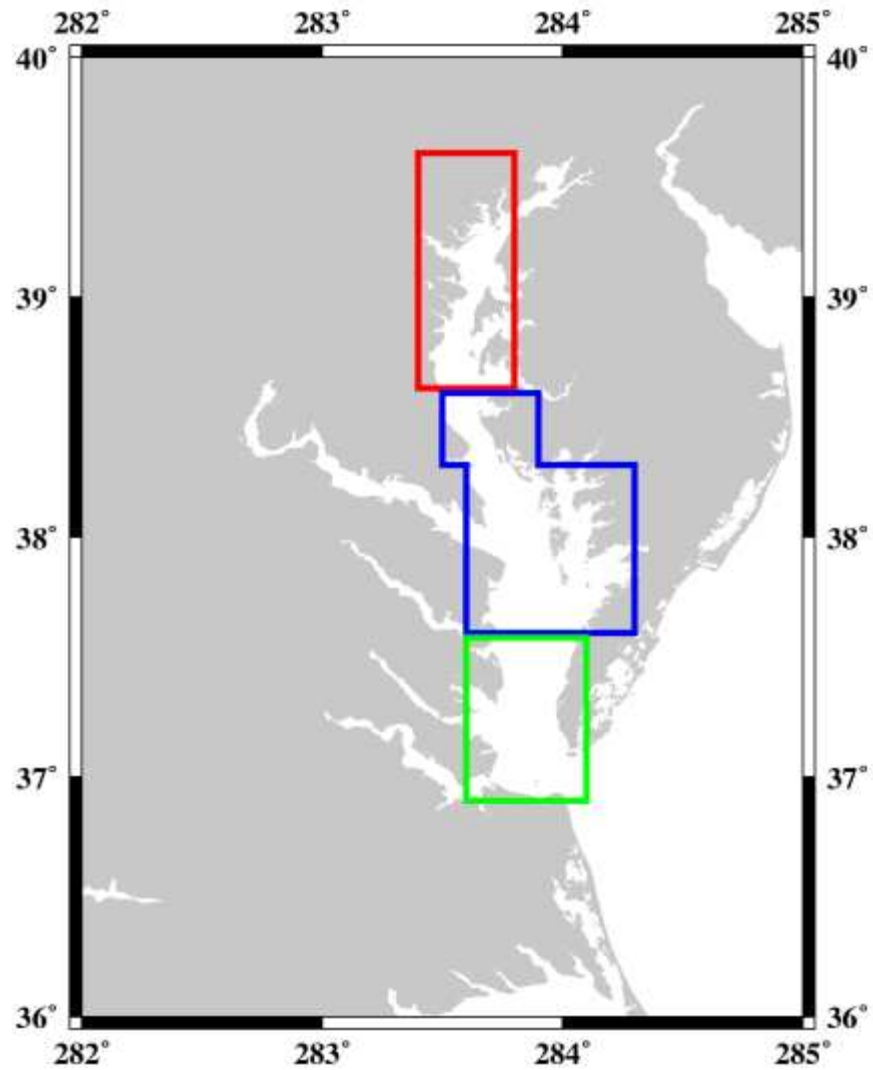


Figure 2. Map of *in situ* C_a data stations. Blue circles indicate the regularly visited Chesapeake Bay Program stations (~15,750 data points). Red circles indicate the Harding stations (~2,300 data points).

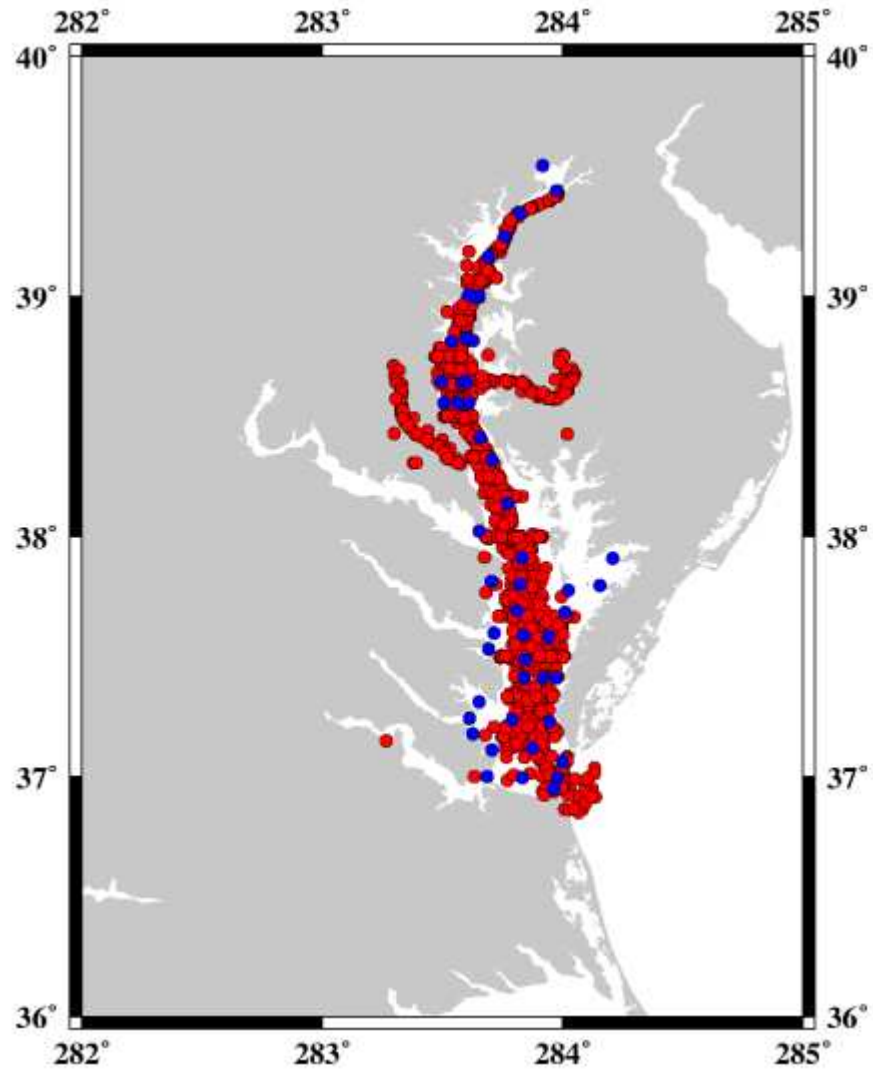


Figure 3. An example of SeaWiFS and *in situ* C_a distributions for the Lower Bay in Spring.

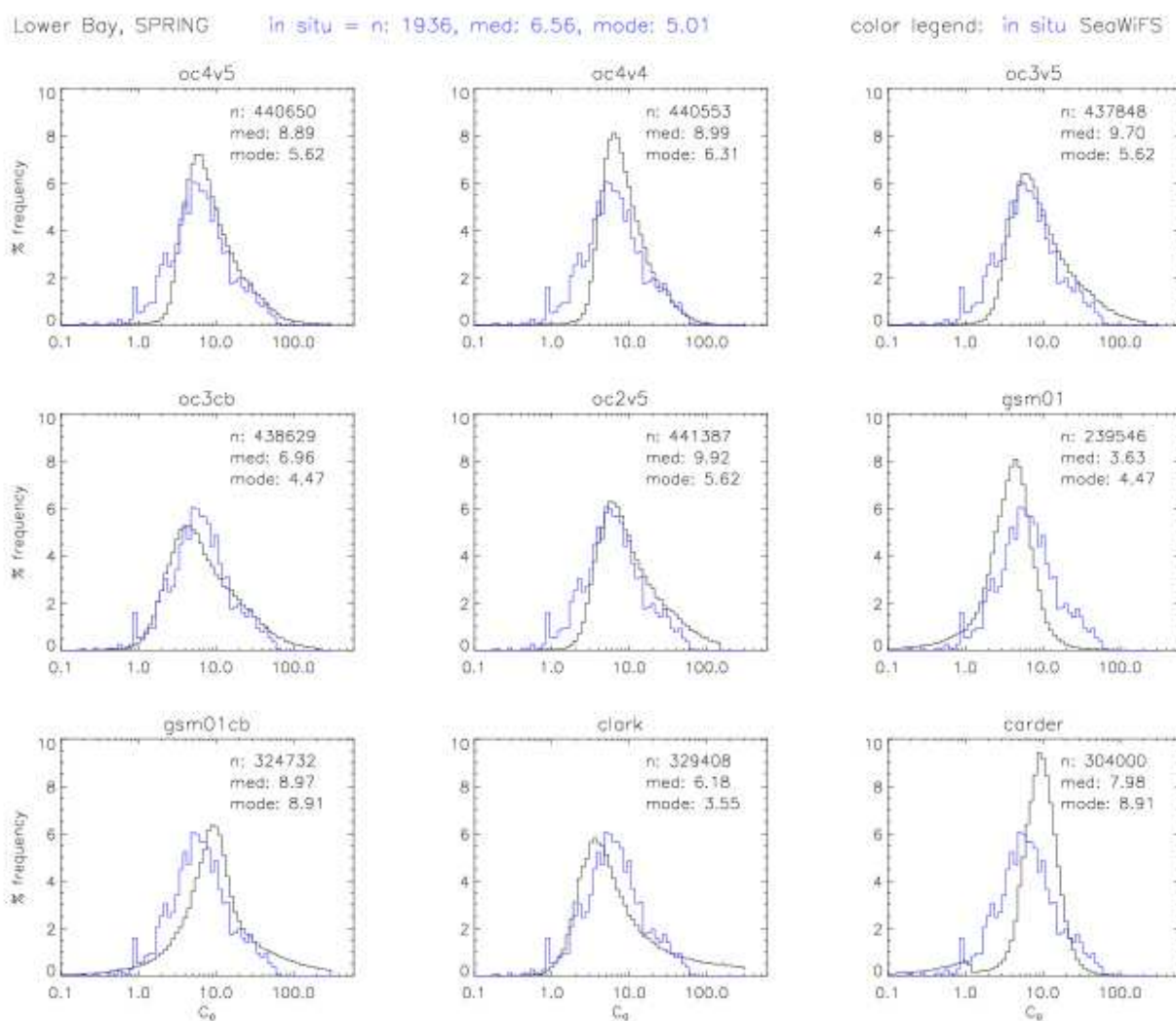


Figure 4. An example of a SeaWiFS and *in situ* C_a time-series for the Lower Bay. For visual reference, the current operational SeaWiFS C_a data product, OC4v4, is included in every plot (blue line). Black circles indicate monthly *in situ* averages. Vertical lines indicate monthly *in situ* standard deviations.

Lower Bay, 1 of 2

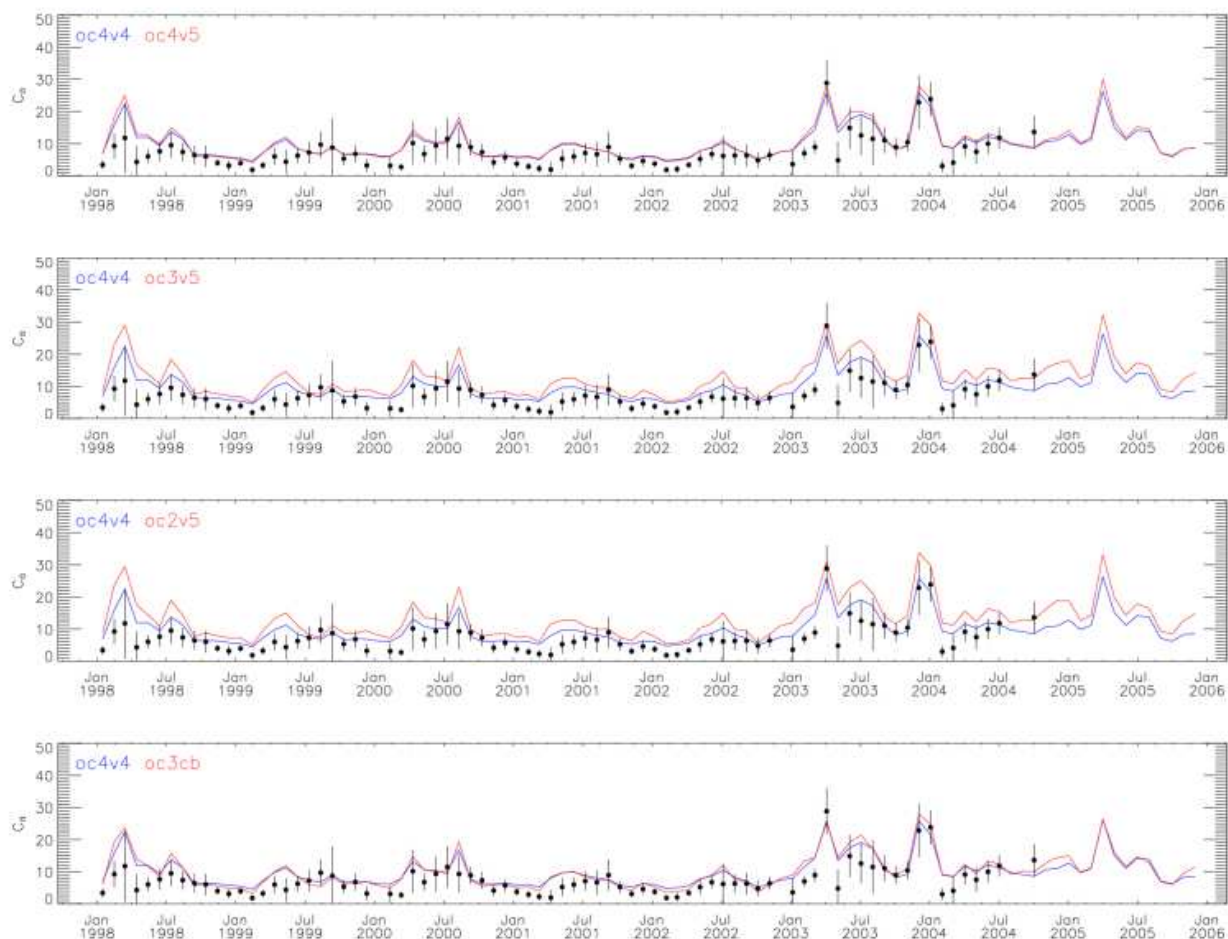


Figure 5. A second example of a SeaWiFS and *in situ* C_a time-series for Algorithm 1 (OC4v4) for the Lower Bay. Black circles indicate monthly *in situ* averages, whereas red circles indicate monthly SeaWiFS averages. Vertical lines indicate monthly standard deviations.

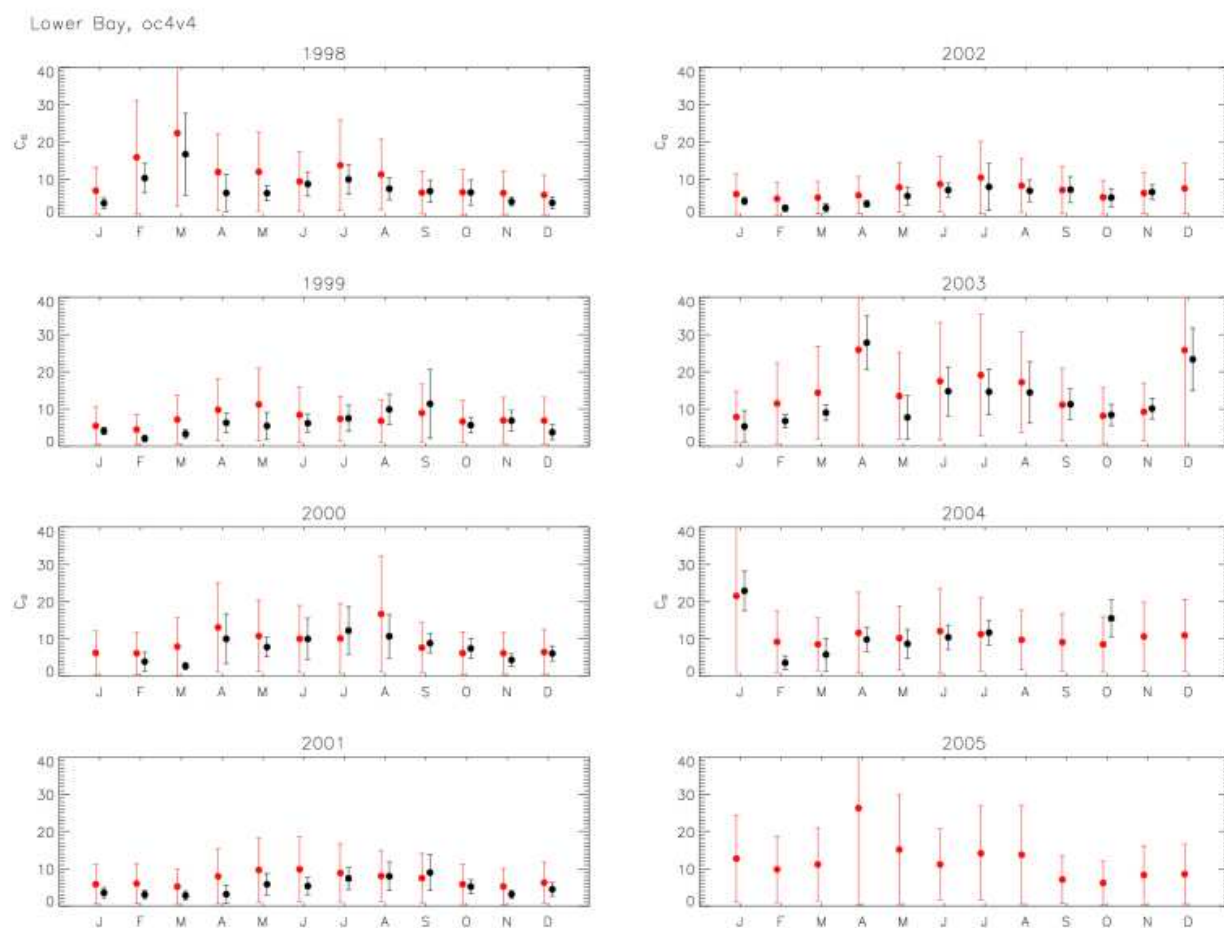


Figure 6. An example of the mapped, extracted true color (left) and OC4v5 C_a (right) browse images. The images were generated from SeaWiFS scene S1997350173253.

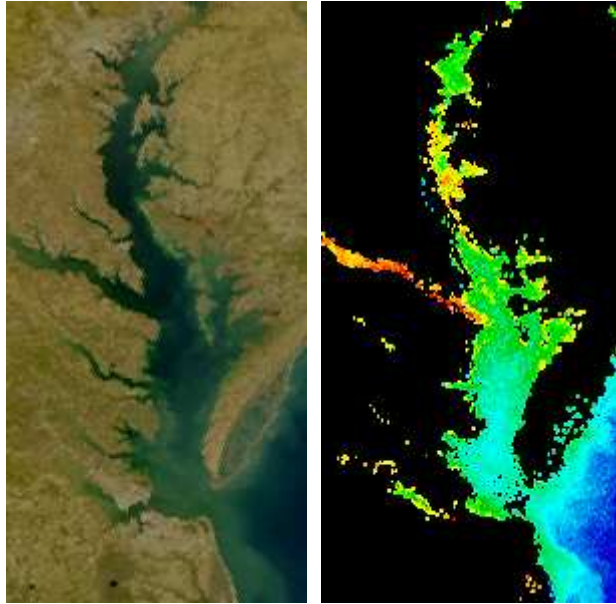


Figure 7. Flowchart of the satellite-to-*in situ* validation process (Figure 1 of Bailey and Werdell (2006)).

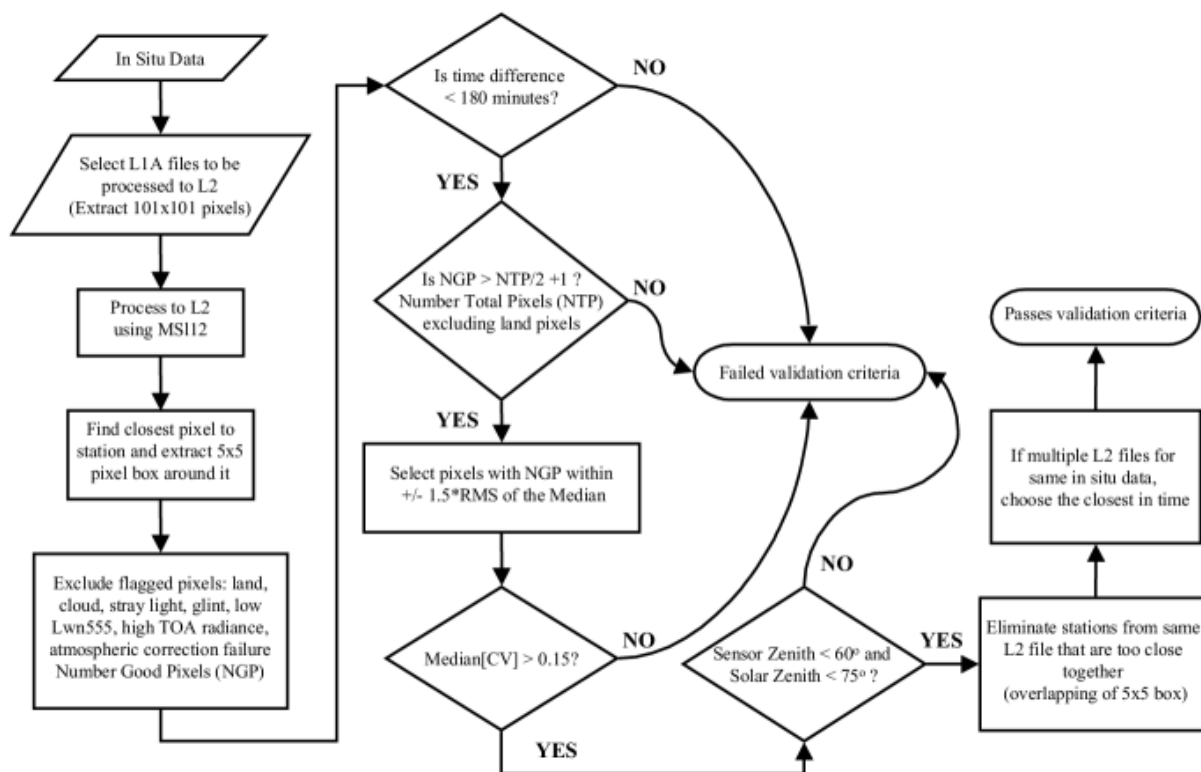


Figure 8. An example of the SeaWiFS versus *in situ* C_a “match-ups” for the Lower Bay in Spring.

